

**KULLANCSINVÁZIÓ • KITERJESZTETT VALÓSÁG • AGYTUMOROK • MAJA KAKAÓ**

LXXVII. évfolyam ■ 42. szám ■ 2022. október 21.

Ára: 500 Ft

Előfizetőknek: 395 Ft

# ÉLET és TUDOMÁNY

## A VULKÁN KÖLELKE



GÉPI TANULÁS SEGÍTÉSE KITERJESZTETT VALÓSÁGGAL

# A JÖVŐ KÉT TECHNOLÓGIÁJÁNAK ÖTVÖZÉSE

**A modern gépi tanulás rohamos iramban fejlődik. Egyre több iparágban egyre szerteágazóbb feladatokra alkalmazzák ezeket a rendszereket. Ugyan a fejlődés folyamatos, egy fontos tényező nem változott a kezdetektől: a megfelelő minőségű, mennyiségű és megfelelően címkézett tanítóadatok rendelkezésre állása nélkül nem lehet jó algoritmusokat fejleszteni. A kutatásom során ennek a problémának a minőségellenőrzésben való enyhítését vizsgáltam kiterjesztett valóság segítségével.**

**M**unkám célja a kiterjesztett valósággal (*augmented reality* -AR) való betanítás egyféle megfordítása volt, ahol nem a felhasználó tanul az eszköz segítségével, hanem egy gépi tanulási algoritmust képez ki. Az elképzelés alapja, hogy az AR eszköz képes egy, a feladatot normálisan végző emberi munkás viselkedése és a környezet elemzése alapján egy automatikusan felcímkézett tanítóadat-halmazt generálni, mely alkalmas egy gépi tanulási rendszer betanítására.

A tanítóadatok hiánya meglehetősen kritikus az Ipar4.0 világában. A kisszerűs gyártás és gyakran változó gyártott elemtípusok miatt sok esetben nincs elég idő megfelelő mennyiségű adat gyűjtésére és nem állnak rendelkezésre létező adatbázisok minden speciális esethez, amelyek segítségével robusztus modellek készülhetnének. A minőségbiztosítás fontos része minden gyártási folyamatnak, azonban a KKV-k többségében ez még mindig szemrevételezéssel, manuálisan történik, a nagyobb cégekkel ellentétben. A gépi tanuláshoz szükséges adathalmaz generálásának megoldása így ezeken a területeken komoly potenciális pénzügyi, biztonsági és egyéb előnyökkel járna az ipari szereplők számára.

Általánosan a gépi tanulás vagy ML (*machine learning*) alatt olyan algoritmusokat értünk, melyek nem explicit programozás által lesznek sikeresek a feladatuk megoldásában. A módszer



A berendezett gyár a virtuális tesztkörnyezetben.

alapja, hogy az algoritmus példa adatok alapján a saját szerkezetét megváltoztatva kerül egyre közelebb egy optimális megoldáshoz. Az egyik legelterjedtebb módja a gép tanulásnak a neurális hálók. Ezek az emberi agy neuronjaihoz hasonló elemekből felépülő hálózatok, melyek a „neuronok” közti csomópontok súlyai alapján hoznak „döntéseket”, és ezen súlyok változtatásával taníthatóak. Kellően nagy méretű hálók képesek a tanító adathalmazban megtanulni a mintákat és azok alapján például képeket osztályokba sorolni, szegmentálni, objektumokat felismerni és sok más tradicionális programozással közel lehetetlen feladatot megoldani.

A tanítóadat halmaz mennyisége és minősége jelentősen befolyásolja az elkészült modell pontosságát, ezért kiemelt hangsúlyt kell fektetni a minőségének biztosítására egy modell tanításakor. Az adathalmaz kiegyensúlyozottsága is fontos szempont, mivel ha a különböző kategóriákról rendelkezésre álló adatok száma nagyban eltér, az jelentősen befolyásolja a modell minőségét. Ez a minőségellenőrzési feladatokban kifejezetten nagy probléma lehet, mivel a modell szempontjából fontosabb „hibás” kategória esetén sokkal nehezebb és időigényesebb megfelelő adathalmazt építeni.

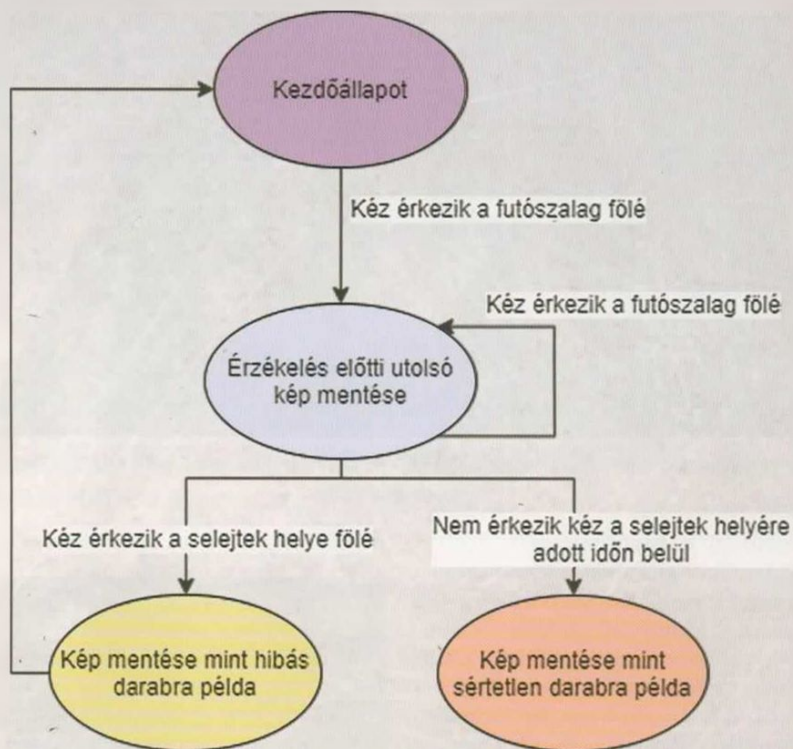
Ilyen algoritmusok tanítása során sok tipikus probléma jelentkezhet, például az overfitting, amikor már nem az



adatban lévő általános mintákat tanulja a modell, hanem a konkrét tanítóadat tulajdonságait veszi egyre jobban figyelembe, így jelentősen rosszabb a hatékonysága valós adatokon. Ennek a problémának több megoldása is van, például a tanítás korábbi megállítása, a tanítóadatban fellelhető zajok szűrése, vagy a tanítóhalmaz kibővítése. Ez utóbbi két pont esetén a jelen kutatás egy értékes lehetőséget nyújthat a probléma felmerülése esetén.

A kiterjesztett valóság olyan technológiákat összefogó gyűjtőfogalom, melyek képesek a valós világra valamilyen módon digitális tartalmakat helyezni. Az egyik legismertebb példája ennek talán a *Pokemon Go* játék, de speciálisabb eszközök sokkal közelebbi hatást képesek elérni a sci-fi filmekben látott hologramokhoz. Az iparban a szemüveg-szerű eszközök terjedtek el leginkább, mivel a kezeket szabadon hagyják a munka végzésére. Ezek fejen viselhető, asztali számítógéptől független eszközök, melyek különleges lencsék segítségével képesek a látótér egy részében virtuális elemeket megjeleníteni a valóság fölött. Ilyen eszköz a *HoloLens 2*, a Microsoft legújabb kiterjesztett valóság szemüvege. A második generációs eszköz megannyi fontos újítást tett elérhetővé az ipari szereplők számára. Az elődjénél jelentősen könnyebb, ezzel segítve a hosszú használatot, a látómező majdnem kétszer olyan nagy részében jelennek meg a hologramok, és számítási kapacitása is jelentősen nőtt. Az első generációban is komoly képességekkel rendelkező szenzorrendszer pedig szemkövető kamerákkal és a kutatás szempontjából kifejezetten fontos kézkövetéssel bővült. Használatuk az iparban egyre szélesebb körű, de mivel még mindig viszonylag drága eszközökről van szó, fontos, hogy minél több területen tudjanak előnyt adni az őket megváltó ipari szereplők számára.

A hipotézisem vizsgálatához prototípust fejlesztettem, melynek megvalósításához több szoftvert is használtam. A *Unity* egy multiplatform általános célú 3D motor, mely széles körű támogatással rendelkezik, mind platformok, fejlesztői eszközök és bővítmények terén. A multiplatform működés kifejezetten fontos szempont



Adatok kiválogatásának módja állapotokkal felírva.

volt, mivel a rendszer a *HoloLens*-en kívül más, esetleg kevesebb lehetőséggel rendelkező eszközön is képes lehet működni, ezzel tovább javítva a költséghatékonyságát, így hasznát az ipari szereplők számára.

A *Mixed Reality Toolkit (MRTK)* egy *Microsoft* által fejlesztett *Unity* bővítmény-csomag, melynek fő célja a *HoloLens* fejlesztés segítése előre elkészített komponensekkel. A keretrendszer fontos felhasználói felületelemeket biztosít, például mutatókat, kézkövetést, kontrollerek kezelését, gombokat, csúszkákat, kapcsolókat és még sok más. Ezen kívül a keretrendszer támogat sok más eszközt is. A projekt szempontjából kifejezetten fontos volt a VR szemüvegek, ezen belül kifejezetten a *Quest 2* támogatása.

Ez azért volt fontos, mert az alkalmazás fejlődése során egy virtuális környezet is elkészült, mely eleinte főleg a fejlesztés gyorsítását szolgálta, azonban később a kísérletet is ebben a környezetben végeztük el. Ezen kívül a gépi tanulásban egyre elterjedtebbek a részben vagy egészen szintetikus adathalmazok. Ezek előállításában nagy segítség lehet egy virtuális környezet. Ipari tervrajzok és látványtervek alapján még

a valódi gyártósor vagy üzem elkészülte előtt elkészíthető egy virtuális mása, melyben virtuális valóság segítségével emberi dolgozók elkezdhetik egy AI tanítóhalmazának elkészítését, mely a szintetikus adatok alapján akár a gyár megnyitására már működőképes is lehet. Ez ugyancsak nagy előny lehet abban az esetben, ha egy emberhez igazított munkaállomás helyett a gyár már rögtön automatizált rendszerek használatához kialakítva épül fel.

A virtuális környezet megjelenítésére a *Quest 2* virtuális valóság szemüveget használtam. Ez az eszköz a valós világot teljesen kizárja, helyette egy teljesen virtuális világot mutat a szemnek. Az eszköz további előnye a nagy felbontása, valamint hogy nem kell számítógéphez kötni, teljesen független működésre is képes. A beépített kézkövetés és számítógéptől független működés miatt az eszköz sok szempontból a *HoloLens 2* virtuális valóságban megfelelője, így ideális választás volt a rendszer tesztelésére, amikor a *HoloLens* nem volt elérhető.

Ezen eszközök és szoftverek segítségével készült el egy *proof-of concept* (a megvalósíthatóságot igazoló – *A szerk.*)





A munkadarabok variációi ahogyan be vannak mutatva a felhasználó számára a felismerés könnyítése céljából.

rendszer a hipotézis tesztelésére. A rendszer a kézfelismerésen kívül a saját térbeli pozícióját is ismeri, így világról alkotott modelljében a valósággal azonos helyen marad a munkadarabokat szállító futószalag és a hibás példányok helye is. Az alkalmazásba érkezik be minden adat és ez generálja a kimeneti adathalmazt. A működés közben az alkalmazás követi a munkás kézmozgását, és az alapján fényképeket készít egy külső kamera segítségével az adott munkadarabról. A rendszer a kézkövetés által biztosított pozíciók alapján címkézi a képeket. Ez úgy lett megvalósítva, hogy a futószalag és selejteknél kijelölt terület fölé a beállítás folyamán egy – egy téglatest kerül, mely figyelni mikor tartalmaz egy kezdet a megadott térfogat.

A sértetlen munkadarabok mentésének feltételei csak akkor teljesülnek, ha a munkás odanyúl egy munkadarabért, de nem veszi fel, vagy visszateszteli miután megvizsgálta. Ha gond lenne az adatok mennyiségével, egy manuális mentés gomb is a munkás rendelkezésére áll, amivel könnyen kibővíthető a helyes elemek halmaza, ha a körülmények megkövetelik.

A rendszer a mentett képeket a címkéknek megfelelően „correct” és „incorrect” mappákba rendezi,

majd menti a HoloLens felhasználó által is hozzáférhető tárhelyére, ahonnan az eszközt számítógépre csatlakoztatva USB-n keresztül lementhetők. Ezek után könnyen el lehet juttatni az adatokat a gépi látás modell tanítását végző program által támogatott formátumban és helyre.

A szimulációs környezetben a kamera virtuális volt, virtuális monitoron keresztül volt elérhető a képe. A valóságban egy IP kamera figyel a futószalagot, melyhez csatlakozik a HoloLens-en futó alkalmazás. Az elhelyezése kritikus és a beszerelés során különös figyelmet érdemel. A fő szempont, hogy előtte legyen annak a pontnak, ahol tipikusan a munkás felemeli a rossz munkadarabokat. A tesztelésre használt virtuális munkadarabokon a sérülések főleg a gyártás során keletkezett ütdések, deformációkat vagy hiányzó elemeket próbálják utánozni.

Az elkészült prototípus hatékonyságát és a továbbfejlesztés irányait egy kísérlet során vizsgáltam. Megfelelő futószalag hiányában ez is virtuális környezetben került lebonyolításra. A kísérlet során az egyik fontos tapasztalat az volt, hogy a virtuális világban sokan nehezen szoknak hozzá a tárgyak megfogásához. Ennek ellenére alapvetően a kísérlet eredményei alapján a virtuális környezet néhány hibáján kívül a rendszer a vártaknak megfelelően teljesít. A felhasználói visszajelzések pozitívak voltak, a rendszer sikeresen valósítja meg a feltételt, miszerint nem zavarhatja a munkást a feladata

elvégzésében. Az általa generált adatok valóban alkalmasak gépi tanulási algoritmus tanítására és azokkal jó hatékonyság érhető el.

Erre a célra az *Inception v3* gépi látási modell egyik példa kódját tanítottam be saját adathalmaz segítségével. Ez a modell a teszteken nyújtott eredményei alapján a legjobb klasszifikálásra tervezett modellek egyike, valamint, a példakód segítségével transzfer tanulás használatával kisebb adathalmaz alapján is nagy hatékonyságú modellt lehet létrehozni vele. Ez a módszer azon az elven alapul, hogy a hatalmas, általános témájú adathalmazokon (pl. *ImageNet*) tanított modelleknek csak az utolsó rétegeit tanítjuk újra a saját céljainkra, így, mivel a nagy adathalmazból a modell már leszűrt általános megkülönböztető jeleket, ezért sokkal kevesebb adattal lehet jó hatékonysági fokot elérni.

A projektet sok irányba lehet az eredmények alapján tovább fejleszteni. Az egyik legfontosabb lépés az ipari tesztekhez egy ipari ellenőrző rendszerbe való integráció lenne. Ezen kívül más, olcsóbb kézkövető eszközökre való átültetése az AR eszközzel nem rendelkező cégek számára magas költségeket csökkentené. Pontosabb felismerést lehetne elérni a HoloLens 2 szenzorrendszer további képességeinek felhasználásával is végül az adatok kinyerését is automatizálni lehetne, az eszköz képes lehet akár egy felhő tárhelyre automatikusan eljuttatni az elkészült képeket, ennek segítségével akár automatizálni is lehet a gépi tanulási modell folyamatos újratanítását.

Úgy gondolom, kutatásom eredménye egy hasznos kiindulópont további fejlesztések, kísérletek elvégzéséhez, amely a piaci szereplőknél bevezetve sok új lehetőséget nyithat a gépi tanulás és a kiterjesztett valóság felhasználásban és terjedésében. A magyar KKV-k hatékonyságának a fejlesztése fontos feladat, valamint a munkaerőhiány, és az egyre magasabb minőségi követelményű autógyártás és hasonló beruházások kiszolgálását is jelentősen segíthetné egy ilyen megoldás.

**SZAKÁL VINCE ABOSA**

BME Automatizálási és Alkalmazott Informatikai Tanszék

Az írás a BME – Pro Progressio  
Alapítvány – Élet és Tudomány 2022. évi  
cikkpályázatára érkezett.

